DOCUMENT RESUME

ED 249 738 EC 170 850

AUTHOR Haring, Norris; And Others
TITLE Investigating the Problem of Skill Generalization: Literature Review I.
INSTITUTION Washington Univ., Seattle. Washington Research Organization.
SPONS AGENCY Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE Jul 84
CONTRACT 300-82-0364
NOTE 59p.; For related document, see EC 170 849.
PUB TYPE Information Analyses (C70)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Basic Skills; Elementary Secondary Education; *Generalization; Literature Reviews; Research Methodology; *Self Evaluation (Individuals); *Severe Disabilities; Training

ABSTRACT

As part of a 5-year investigation of skill generalization in severely handicapped students, a literature review was undertaken within each of three general approaches: performance pattern studies, ecological studies, and self-control studies. The first paper, "Descriptive Analysis of Extant Research Literature concerning Skill Generalization and the Severely/Profoundly Handicapped" by Owen R. White, reviews 30 studies from the Washington Research Organization (UWRO) and concludes that only 13 of the studies provided information useful in systematically refining an instructional technology for promoting skill generalization with severely handicapped persons. Ann K. Berman and Cheryl L. Opalski ("The Impact of Functional Trial Sequencing upon Generalization") review effects on skill generalization of the nature of the task, the environment, and the task's impact. In the final paper, "Self-Monitoring and Skill Generalization: A Review of Current Research," Kathleen A. Liberty focuses on studies of self-monitoring, one of three most common self-control techniques. (CL)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

4242

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Norris Haring

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

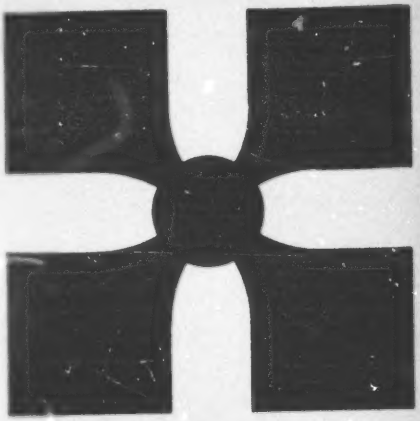
ED249738

Washington Research Organization

Research in Education of the Severely Handicapped
University of Washington

# INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION:

# LITERATURE REVIEW I





EC170850

Washington
Research
Organization

Research in Education of
the Severely Handicapped

University of Washington
205A Parrington Hall, DC-05
Seattle, WA 98195

# INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION:

# LITERATURE REVIEW I

EC170850

Norris Haring, Principal Investigator
Kathleen Liberty, Project Coordinator

Michael Boer, Editor

The activity which is the subject of this report was supported in whole or in part by the U. S. Department of Education (Contract No. 300-82-0364). However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Printed in the United States of America.

July 1984

## Preface

The Washington Research Organization (UWRO), an Institute for Research in Education of the Severely Handicapped, is conducting a five-year investigation of the problem of skill generalization. During its first project year, UWRO investigated three approaches to developing strategies for facilitating skill generalization:

(1) *Performance Pattern Studies* began with a retrospective analysis of existing data sets, and will proceed to the collection of descriptive data in public school classrooms. These data will be used to determine a set of experimental decision rules for matching specific instructional methods to individual learners.

(2) *Ecological Studies* initiated a four-year longitudinal descriptive study of factors in educational settings which may influence generalization. Intervention studies are also included within this approach, starting with a pilot study of massed vs. distributed instructional trials on generalization.

(3) *Self-Control Studies* commenced with three studies of the effects of self-monitoring procedures on skill generalization. This area of studies will also include investigations of self-reinforcement and self-instruction.

During the first year of research, UWRO investigators have conducted reviews of applicable literature within each of these general approaches:

Dr. Owen R. White presents, in "Descriptive Analysis of Extant Research Literature Concerning Skill Generaliza-

tion and the Severely/Profoundly Handicapped," a descriptive analysis of 30 published studies from UWRO's retrospective analysis of existing data sets.

Ann K. Berman and Cheryl L. Opalski's "The Impact of Functional Trial Sequencing Upon Generalization" provides a review of the literature that pertains to one aspect of UWRO's ecological studies.

Dr. Kathleen A. Liberty's review of the literature, "Self-Monitoring and Skill Generalization: A Review of Current Research," is related to the subject of her first year of UWRO studies in self-control.

UWRO will produce a total of four literature review products, of which this is the first. Each of the current research approaches will continue through the next four years, with the ultimate objective of producing a unified set of guidelines for practitioners to use in facilitating skill generalization.

**Norris G. Haring**
**Principal Investigator**
**Seattle, 1984**

## Table of Contents

**Preface** v

**Table of Contents** vii

**White, Owen R.** "Descriptive analysis of extant research literature concerning skill generalization and the severely/profoundly handicapped." 1

**Berman, Ann K. and Opalski, Cheryl L.** "The impact of functional trial sequencing upon generalization." 21

**Liberty, Kathleen A.** "Self-monitoring and skill generalization: A review of current research." 37

# DESCRIPTIVE ANALYSIS OF EXTANT RESEARCH LITERATURE CONCERNING SKILL GENERALIZATION AND THE SEVERELY/PROFOUNDLY HANDICAPPED

**Owen R. White**

The following analyses are based on 30 studies of skill generalization and severely/profoundly handicapped persons. The general procedures for this retrospective analysis of published data included the development of a form for coding information about the studies, development of coder reliability on the content and format of the coding form, entry of coded information in a computer, computerized summaries of the data collected, and analysis.

Articles were selected for coding on the basis of the following criteria: (1) Data on generalization must be included in the article. (2) The data must show responses of at least one individual subject (group data alone are not sufficient). (3) The study must concern at least one behavioral target that is an acceleration target. (4) The study must have been published since Stokes and Baer's (1977) article summarizing generalization studies to date.

In addition, since there were a wide variety of professional journals from which to choose, journals that had a high probability of containing studies that met our criteria were searched first (e.g., *Journal of Applied Behavior Analysis* was searched before *American Journal of Mental Deficiency*). An effort was made to include articles where at least one of the subjects was severely, profoundly, or multiply handicapped, but other studies were also included, since studies with more severely handicapped subjects that met our criteria were difficult to identify. Overall, median reliability for article coding was 88%.

The analyses provided below are descriptive in nature, providing information concerning the frequency with which each coded condition occurred in the articles reviewed. Observed conditions are evaluated in terms of their implications for the study of generalization *per se* and the utility of available data in resolving certain questions being posed by UWRO. No attempt is made to draw conclusions concerning the relationships among variables (e.g., the relationship between various types of generalization strategies and eventual success) due to the limited number of studies thus far reviewed and coded. Certain relationships do appear to be emerging, however, and will be reported if confirmed by further review.

## Subjects

### Number of Subjects

The number of subjects in each study or series of related studies was generally small. The actual range of subject sample sizes extended to a maximum of 12, but a boxplot analysis (Tukey, 1977) revealed that the typical distribution ended at a maximum of 8 subjects per study, and that the 3 studies employing 12 subjects should be considered exceptions to the general case.

As all studies reviewed employed some form of single-subject analytic strategy, the small sample sizes *per se* did not necessarily threaten internal validity. Given the types of handicapping conditions represented by that small number of subjects, however, the external validity of encoded studies with respect to the population of "severely handicapped persons" as a whole must be questioned.

Table 1-1

**Disabilities Represented by Subjects**

| Disability Type | Number of Studies |
|---|---|
| Normal (no disabilities) | ☑☑☑☑ |
| Mild Mental Retardation | ☑ |
| Moderate Mental Retardation | ☑☑☑☑☑☑☑☑☑☑☑ |
| Severe Mental Retardation | ☑☑☑☑☑☑☑☑☑ |
| Profound Mental Retardation | ☑☑☑☑☑☑☑☑☑ |
| Multiply Handicapped | ☑☑ |
| Physically Handicapped | ☑ |
| Visually Impaired | |
| Auditorily Impaired | |
| Communication Impaired | |
| Behavior Disability | ☑ |
| Deaf-Blind | |
| Other Handicapping Condition | ☑ |

**Subject Types**

The distribution of disability types represented in the 30 studies is illustrated in Table 1-1. The total number of cases noted (38) exceeds the number of actual studies reviewed (30) due to cases in which studies involved subjects with more than one disability type.

All studies reviewed involved at least one subject who could be considered severely handicapped in some manner. Disability types which could not in themselves constitute a severe handicap occurred in studies employing control or contrast subjects for other, more severely handicapped subjects.

UWRO reviewers attempted to concentrate on studies dealing with the severely mentally retarded. It is not possible, therefore, to make general statements concerning the representation of various disability types in the entire body of research literature as a whole. Given that a high proportion of severely mentally retarded persons display one or more other handicaps as well, however, the poor representation of subjects with physical, visual, and auditory handicaps must be considered a serious threat to the overall external validity of the research reviewed to date.

**Subject Age**

Subjects ranged in age from less than 2 years to in excess of 50 years. The majority of studies involved subjects ranging from circa 6 years of age to 21 years of age (67% of the studies). Generally, therefore, the studies reviewed appeared to provide adequate representation of the school population with respect to age.

**Summary**

Since an attempt was made to review articles studying skill generalization with severely mentally retarded subjects, no statements can be made concerning the degree to which all available research represents the total population of severely/profoundly handicapped. Within that constraint, however, most studies deal with a small number of subjects (6 or less) in the 6-to-21 year age range. Notably, no studies were reviewed which included subjects with physical, visual, or auditory handicaps, so for the most part, conclusions based upon the UWRO review must be limited to what might be called a sampling of "physically intact" mentally retarded individuals.

## Behaviors Studied

### Behavior Class

Most behaviors studied fell into one of six major classifications: social, communication, vocational/prevocational, self-help/independent living, cognitive/academic skills, and cognitive strategies. The majority of studies investigated the development of communication skills (e.g., expression of need), self-help or independent living skills (e.g., use of local transportation systems), social skills (e.g., engaging in cooperative play behaviors), or academic/conceptual behavior (e.g., the discrimination of "survival" words in the environment). Somewhat surprisingly, 13 percent of the studies investigated general cognitive strategies (e.g., strategies for eliciting reinforcement from other persons in the environment) while only circa 10% of the studies investigated vocational or prevocational skills (e.g., janitorial skills).

### Functionality of Behavior

The great majority (93%) of behaviors selected for study were perceived by both the original investigators and the reviewers as being of immediate functional utility to the subject. That is in conformance with recent instructional trends, and should facilitate generalization by increasing the probability of natural consequation outside the instructional situation.

### Complexity of the Behavior

Training targets varied considerably in complexity, ranging from simple one-step behaviors (e.g., "responds appropriately to yes-no questions") to tasks in which 180 separate subtasks were identified (e.g., "cleans bathroom"). However, nearly two-thirds of the investigated tasks contained fewer than ten separate subtasks.

### Pretraining Performance Level

Most studies (74%) involved the development of new skills where pretraining performance levels were demonstrated or assumed to be zero. At times the study of generalization of such skills was incidental to the demonstration of skill acquisition, although the development of new skills was most commonly a vehicle for controlled studies of eventual skill generalization.

Occasionally investigators continued to refine skills after initial acquisition or work with behaviors that had already been acquired at some level in an attempt to produce generalized responding, or to work with a behavior which had already met "criterion" in one setting or situation but which had not yet generalized appropriately.

The relatively low incidence of studies involved in the systematic refinement of existing skills is a disappointment to UWRO investigators. It is hypothesized that the fluency with which a person is able to engage in a behavior during instruction will be related to the probability of eventual generalization, so it is unfortunate that few investigators studied skill refinement or, indeed, even collected the type of performance data which would enable the issue of fluency to be studied at all (see the section on "Assessment Data," below).

## General Settings and Conditions During Initial Training

### Settings

Many studies involved initial training within public school settings (37%), but the most common training setting was a special laboratory or room (47%), albeit, that room may have been located within a public school. Very few studies conducted initial training within the community or home (10% each). A higher percentage (50%) did study generalization within the "natural setting" after initial training (see "Generalization Conditions and Strategies," below), but due to the nature of the skills being studied, those natural settings tended to be school, special vocational, or residential facility environments, rather than the community or home. That would seem contrary to the current emphasis on community integration and school-home cooperative programs.

### Trainer Type

The majority of studies (77%) employed a member of the research staff (experimenter or research assistant) to implement initial skill training programs. Only 10% of the studies employed the subject's regular teacher, therapist, or counselor

during initial skill training, and only circa 7% employed the subject's parent or guardian. Dispite the apparent simplicity of most instructional procedures employed, therefore, little direct information exists in the reviewed literature concerning the ability of regular practitioners or parents to implement the procedures under study.

**Trainer/Subject Ratio**

All but two studies conducted initial training sessions in a one-to-one setting (one trainer with one subject). Only one study employed a group situation (one trainer to more than five subjects), and one study employed several peer trainers with individual handicapped subjects. While effective small or large group training alternatives would be desirable, most training in classes for the severely handicapped is still conducted on a one-to-one basis, so the use of such procedures in the studies reviewed cannot be considered a serious impediment to eventual adoption within applied settings.

**Antecedents/Cues/Prompts**

A wide range of instructional cues and prompts were used during initial skill training. Most, however, were "task specific," designed primarily to elicit a particular response. Only circa 23% of the cues and prompts could be considered specifically designed to promote a form of generalized response class (e.g., generalized imitation in response to the cue, "do this"). Moreover, most cues were "artificial" (i.e., would not occur in the natural environment), with only 3 studies employing some mixture of what might be considered "natural" and "artifical" cues, and only a single study employing what might be considered "all natural" cues (i.e., cues which could occur in the natural environment).

**Consequences/Reinforcement**

Most studies employed some mixture of social consequation, correction procedures, and/or repeated trials (for errors) during initial skill development. More studies used at least a mixture of natural and artificial consequences than was the case with cues or prompts, but only a single study used what might be considered "all natural" consequences during initial skill training.

**Arrangement/Contingency/Schedule**

Approximately two-thirds of the studies employed continuous schedules of consequation (1:1) during some part of initial training. Such contingencies are not noted for their resistance to extinction and would, for the most part, reduce the probability of successful generalization. One-third of the studies did eventually employ some systematic method for adjusting or fading the frequency of consequation, but somewhat more studies employed fixed schedules of consequation rather than variable schedules (the latter would generally be considered more resistant to extinction). Finally, only one study attempted to base consequation (at least in part) on a temporal feature of the behavior (latency), and in that study the subject was allowed 10 seconds to respond before correction procedures were undertaken. It would seem, therefore, that fluency of response in most studies was not considered of sufficient concern to warrant contingencies which might explicitly facilitate its development.

**Initial Training Effectiveness**

Since each reviewed study used idiosyncratic training procedures, and few studies conducted direct comparative analyses concerning different approaches to training, virtually nothing can be said concerning the relative efficacy of initial training procedures *per se*. Unlike the literature of a decade ago, however, the great majority (87%) of studies were able to teach all subjects the initial training task, and all but one of the subjects who did not complete the initial training procedure were dropped due to conditions beyond the control of the investigators (e.g., subject illness or transfer away from the training site). It might seem that skill acquisition is less of a problem than it once was, and that a focus on the development of procedures which facilitate generalization after initial skill acquisition is appropriate. One should be careful in accepting that conclusion, however, inasmuch as many studies carefully screened subjects before training, at least in part to ensure a subject sample that would be likely to acquire the task in question.

## Generalization Conditions and Strategies

### Strategies Used to Promote Generalization

Virtually all of the training strategies identified by Stokes and Baer (1977) which might lead to generalization were represented in the studies reviewed (see Table 1-2). As found by Stokes and Baer, the most common "strategy" (actually, a lack of strategy) was the simple "train and hope" model. A somewhat greater proportion of the recent literature employed "loose training," "sufficient exemplars," and "natural maintaining contingencies" than Stokes and Baer observed, but no studies were reviewed that employed the use of "indiscriminable contingencies" and only two studies used procedures which involved specific contingencies designed to foster the development of skill variation/generalization *per se* (i.e., "train to generalize"). One additional strategy, not originally mentioned by Stokes and Baer, appears to have gained considerable popularity: "training in the natural environment." A few studies conducted comparative analyses of at least two strategies, but in virtually all of those studies the comparison was limited to "train and hope" versus some more formal procedure for promoting generalization. Very little information is available, therefore, concerning the relative efficacy of alternative strategies specifically designed to promote generalization. Also, for obvious reasons, the "train and hope" procedure was always attempted first, so the effect of the alternative procedure was always confounded with intervention-sequence.

### Changes in Conditions During Generalization Probes

Collectively, reviewed studies probed for generalization across a wide range of conditions (see Table 1-3). Although some studies involved only what might be considered "trivial" generalization parameters (e.g., generalization across time of day or specific manager), most made some effort to probe for generalization within the setting or situation where the trained skill would most naturally be employed. In four studies (13%), response requirements under generalization probe conditions were also changed, requiring some form of response adaptation or induction in addition to simple stimulus generalization.

Table 1-2

**Generalization Training Strategy**

| Strategy | Number of Studies |
|---|---|
| Train & Hope | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Natural Setting | ☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Sequential Modification | ☑☑ |
| Train Loosely | ☑☑☑☑☑ |
| Sufficient Exemplars | ☑☑☑☑☑☑☑ |
| Common Stimuli | ☑☑ |
| Natural (aint. Contingencies | ☑☑☑☑☑☑ |
| Mediate Generalization | ☑ |
| Train to Generalize | ☑☑ |

Table 1-3

**Changes iz Conditions during Generalization Probes**

| Condition Change | Number of Studies |
|---|---|
| Settings | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Time of Day | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Items/Stimuli | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| People | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Ant./Cues | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |
| Consequences | ☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑☑ |

## Assessment Data

### Training Data

Training data were provided in 26 (87%) of the studies reviewed; four studies reported only that "training was conducted successfully." With those studies, of course, no analysis of the relationship between initial acquisition curves and eventual generalization is possible. All of the 26 studies reporting training data did report the data for each individual subject, but in only six (20%) of the cases was it possible to identify precisely how many responses (correct/error) were made during each training session. The remaining studies reported data in "blocks of sessions" or as some computed statistic (e.g., "percent of possible," where the number possible was not reported). Such summaries can seriously distort the actual nature of an acquisition curve and obviate meaningful analysis of its relationship to eventual skill generalization. Finally, only two (7%) of the studies reported time-based data (e.g., rate, latency) for initial training, obviating the possibility of evaluating the relationship between initial response fluency and eventual generalization. The lack of time-based data during initial training also limits the evaluation of changes in response characteristics during generalization probes. That is, while it might be possible to determine whether subjects did or did not generalize at all on a given occasion, it was not possible with most studies to determine whether there was any decrease in response fluency under generalization conditions.

### Generalization Probes

In order to form a complete picture of the relationship between intial training and eventual skill generalization, systematic probes for generalization should probably occur before, during, and after each phase of skill training. Unfortunately, that was not always the case.

Seven studies (23%) did not probe for skill generalization prior to the initiation of training. In most of those cases baseline training probes (i.e., probes to determine pretraining level of performance in the training situation) indicated less than criterion performance, so one might assume that no satisfactory generalized responding would occur.

Six studies (20%) did not probe for generalization at intervals during the initial training process. With those studies it would

not be possible to determine whether meaningful generalization occurred prior to the point when the subjects reached the often arbitrary criterion for initial skill acquisition.

Eight studies (27%) failed to probe for generalization after the termination of training. Such studies did probe for generalization at least once near or at the end of training, but would have no measure of "maintenance" of effect, even over very brief periods. Five studies (17%) probed for generalized responding *only* after training had been concluded. In those studies, no assurances are provided that generalized responding did not occur prior to training, and no estimate of when generalized responding might first have occurred (or been possible, given the opportunity) was possible.

It should be noted that there are potential justifications for minimizing the number of generalization probes undertaken during a study. Aside from the potential difficulty and cost of such probes, frequent generalization probes may constitute one or more confounds. For example, if reinforcement (natural or otherwise) is provided for appropriate responding in generalization situations, that reinforcement would constitute a potential "training strategy" for continued responding, obviating any "clean" assessment of further generalization *per se*. Conversely, if reinforcement is withheld in generalization probes, repeated exposure to those conditions would constitute "extinction conditions" and could lead to the formation of a discrimination between training and nontraining situations, the opposite of what one actually hopes to accomplish. It is necessary, therefore, to consider the issue of repeated and frequent generalization probes very carefully, but the failure to probe for generalization before, during, and after initial training should be avoided whenever possible.

## Outcomes and Implications

### Level of Generalization

In all but three of the 30 studies reviewed, authors characterized their subjects as generalizing "some" or "well." However, authors characterized *all* of the subjects in any given study as generalizing "well" in only 11 (37%) of the cases. UWRO reviewers tended to be somewhat more conservative, recording that in their opinion, only 2 (7%) of the studies demonstrated

"complete generalization," and that only 12 (40%) of the studies appeared to employ methods which facilitated some meaningful generalization in most subjects. The discrepancy in opinions might be due to an unconscious predisposition to find success on the part of the authors, but in at least some cases authors were actually more reserved in their opinions than the reviewers. More likely, therefore, the discrepancy is simply due to a lack of concensus concerning the level of performance that constitutes "meaningful generalization."

**Applicability of Findings**

Reviewers felt that 16 (53%) of the studies employed procedures that could most likely be applied in a wide range of nonresearch settings. In the remaining cases, however, one or more features of the method(s) employed would likely serve as a serious impediment to its widespread adoption. In most of those cases, the major problem appeared to be a simple matter of cost, either because of the personnel and time required to implement the procedure, or the special equipment/settings involved.

**Replicability of Studies**

Surprisingly, considering that all reviewed studies were published in research journals, reviewers only felt that 21 (70%) of the studies were described in sufficient detail to allow precise replication. Problems centered most often on descriptions of training procedures or the conditions under which generalization was assessed. For example, one study described training procedures as "modeling, role-playing, and social consequation." It should be noted, however, that precise replication of all procedures might not be required in order to achieve a functional replication of outcomes, and additional information concerning procedures is often available from authors upon request.

**Implications for Development of a Technology of Generalization Instruction**

The demonstration that a particular procedure is capable of facilitating skill generalization does not necessarily further the development of a precise instructional technology. In order for such a technology to evolve, studies must be conducted which systematically analyze variation in approach in a manner which allows specific controlling variables to be identified. Unfortunately, most of the reviewed studies did not lend themselves to

such analyses. As mentioned earlier, most studies merely compared a given approach with the "train and hope" paradigm, demonstrating at best that some directed attempt to facilitate generalization was better than no formal attempt. In those studies which did make some attempt to compare alternative strategies for the active promotion of generalization, the "favored" strategy almost always followed an attempt to produce generalization with less favored strategies -- raising the possibility of an intervention sequence confound. Overall, reviewers were of the opinion that only 13 (43%) of the studies provided information useful in systematically refining an instructional technology for promoting skill generalization with severely handicapped persons; and indeed, many of those studies still contained potentially serious experimental confounds. The remaining 17 (57%) of the studies may have provided an example of how one *might* facilitate skill generalization, but did not provide sufficient information to enable the scientist to refine hypotheses concerning controlling variables, or the practitioner to make an informed choice among viable instructional alternatives.

## References


Anderson, S., & Spradlin, J. (1980). The generalized effects of productive labeling training involving common object classes. *Journal of the Association for the Severely Handicapped, 5*, 143-157.

Campbell, C., & Stremel-Campbell, K. (1982). Programming "loose training" as a strategy to facilitate language generalization. *Journal of Applied Behavior Analysis, 15*, 295-301.

Coon, M., Vogelsberg, R., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for the Severely Handicapped, 6*, 46-53.

Halle, J., Marshall, A., Spradlin, J. (1979). Time delay: A technique to increase language use and facilitate generalization in retarded children. *Journal of Applied Behavior Analysis, 12*, 431-439.

Stimbert, V., Minor, J., & McCoy, J. (1977). Intensive feeding training with retarded children. *Behavior Modification, 1*, 517-529.

Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Stokes, T., Fowler, S., & Baer, D. (1978). Training preschool children to recruit natural communities of reinforcement. *Journal of Applied Behavior Analysis*, 11, 285-303.

Thompson, T., Braam, S., & Fuqua, W. (1982). Training and generalization of laundry skills: A multiple probe evaluation with handicapped persons. *Journal of Applied Behavior Analysis*, 15, 177-182.

Tucker, D., & Berry, G. (1980). Teaching severely retarded multihandicapped students to put on their own hearing aids. *Journal of Applied Behavior Analysis*, 13, 65-75.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, PA: Addison-Wesley.

van den Pol, R., Iwata, B., Ivancil, M., Page, T., Neef, N., & Whitley, F. (1981). Teaching the handicapped to eat in public places: Acquisition, generalization, & maintenance of restaurant skills. *Journal of Applied Behavior Analysis*, 14, 61-69.

Welch, S., & Pear, J. (1980). Generalization of naming responses to objects in the natural environment as a function of training stimulus modality with retarded children. *Journal of Applied Behavior Analysis*, 13, 629-643.

**Bibliography of Articles Coded**

Adkins, J., & Matson, J. L. (1980). Teaching institutionalized mentally retarded adults socially appropriate leisure skills. *Mental Retardation*, 18, 249-252.

Bornstein, P. H., Bach, P. J., McFall, M. E., Friman, P. C., & Lyons, P. D. (1980). Application of a social skills training program in the modification of interpersonal deficits among retarded adults: A clinical replication. *Journal of Applied Behavior Analysis*, 13, 171-176.

Bornstein, M. R., Bellack, A. S., & Hersen, M. (1977). Social-skills training for unassertive children: A multiple-baseline analysis. *Journal of Applied Behavior Analysis*, 10, 183-195.

Browder, D. M., Moris, W. W., Snell, M. E. (1981). Using time delay to teach manual signs to a severely mentally retarded student. *Education and Training of the Mentally Retarded*, 16, 252-258.

Burgio, L. D., Whitman, T. L., Johnson, M. R. (1980). A self-instructional package for increasing attending behavior in educable mentally retarded children. *Journal of Applied Behavior Analysis*, 13, 443-460.

Burney, J. D., Russel, B., Shores, R.E. (1977). Developing social responses in two profoundly retarded children. *American Association for the Education of the Severely/Profoundly Handicapped Review*, 2, 53-60.

Campbell, C. R., & Stremel-Campbell, K. (1982). Programming "loose training" as a strategy to facilitate language generalization. *Journal of Applied Behavior Analysis*, 15, 295-302.

Combs, M. L. & Lahey, B. (1981). A cognitive social skills training program. *Behavior Modification*, 5, 39-59.

Cooke, T. P., & Apolloni, T. (1976). Developing positive social-emotional behaviors: A study of training and generalization effects. *Journal of Applied Behavior Analysis*, 9, 65-78.

Coon, M. E., Vogelsberg, R. T., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for the Severely Handicapped*, 6, 46-53.

Cuvo, A. J., Leaf, R. B., & Borakove, L. S. (1978). Teaching janitorial skills to the mentally retarded: Acquisition, generalization and maintenance. *Journal of Applied Behavior Analysis*, 11, 345-355.

Dineen, J. P., Clark, H. B., & Risely, T. R. (1977). Peer tutoring among elementary students: Educational benefits to the tutor. *Journal of Applied Behavior Analysis*, 20, 231-238.

Dowrick, P. W., & Dove, C. (1980). The use of self-modeling to improve the swimming performance of spina bifida children. *Journal of Applied Behavior Analysis*, 13, 51-56.

Farb, J., & Thorne, J. M. (1978). Improving the generalized mnemonic performance of a Down's syndrome child. *Journal of Applied Behavior Analysis*, 11, 413-419.

Filler, J., & Kasari, C. (1981). Acquisition, maintenance & generalization of parent-taught skills with two severely handicapped infants. *Journal of the Association for the Severely Handicapped*, 6, 30-38.

Giangreco, M. F. (1983). Teaching basic photography skills to a severely handicapped young adult. *Journal of the Association for the Severely Handicapped*, 8, 43-49.

Hall, C., Sheldon-Wildgen, J., & Sherman, J. A. (1980). Teaching job interview skills to retarded clients. *Journal of Applied Behavior Analysis*, 13, 433-442.

Hart, B. & Risley, T. R. (1980). In vivo language intervention. *Journal of Applied Behavior Analysis*, 13, 407-432.

Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behavior in severely handicapped youth: Pinball machine use. *Journal of the Association for the Severely Handicapped*, 6, 38-44.

Hill, J. W., Wehman, P., Pentecost, J. (1980). Developing job interview skills in mentally retarded adults. *Education and Training of the Mentally Retarded*, 15, 179-186.

Hodges, P. M., & Deich, R. F. (1978). Teaching an artificial language to non-verbal retardates. *Behavior Modification*, 2, 489-509.

Jackson, G. M. (1979). The use of visual orientation feedback to facilitate attention and task performance. *Mental Retardation*, 6, 281-284.

Johnson, B. F., & Cuvo, A. J. (1981). Training mentally retarded adults to cook. *Behavior Modification*, 5, 187-202.

Kissel, R. C., Johnson, M. R., & Whitman, T. L. (1980). Training a retarded client's mother and teacher through sequenced instructions to establish self-feeding. *Journal of the Association for the Severely Handicapped*, 5, 382-394.

Kissel, R. C., & Whitman, T. L. (1977). An examination of the direct and generalized effects of a play-training and overcorrection procedure upon the self-stimulatory behavior of a profoundly retarded boy. *American Association for the Severely and Profoundly Handicapped Review*, 2, 131-146.

Koegel, R. L., Egel, A. L., & Williams, J. A. (1980). Behavioral contrast and generalization across settings in the treatment of autistic children. *Journal of Experimental Child Psychology*, 30, 422-437.

Kohl, F. L., Karlan, G. R., & Heal, L. W. (1979). Effects of pairing manual signs with verbal cues upon the acquisition of instruction following behaviors and the generalization to expressive language with severely handicapped students. *American Association for the Severely and Profoundly Handicapped Review*, 4, 291-300.

Lahey, B. B., Busmeyer, M. K., O'Hara, C., & Beggs, V. E. (1977). Treatment of severe perceptual-motor disorders in children diagnosed as learning disabled. *Behavior Modification*, 1, 123-140.

Lancioni, G. . & Ceccarani, P. S. (1981). Teaching independent toileting within the normal daily program: Two studies with profoundly retarded children. *Behavior Research of Severe Developmental Disabilities*, 2, 79-96.

Matson, J. L. (1981). Assessment and treatment of clinical fears in mentally retarded children. *Journal of Applied Behavior Analysis*, 14, 287-294.

Matson, J. L., & Adkins, J. (1980). A self-instructional social skills training program for mentally retarded persons. *Mental Retardation*, 10, 245-248.

Matson, J. L., & Andrasik, F. (1982). Training leisure time social-interaction skills to mentally retarded adults. *American Journal of Mental Deficiency*, 86, 533-542.

Miller, S. M. & Sloane, H. N. (1976). The generalization effects of parent training across stimulus settings. *Journal of Applied Behavior Analysis*, 9, 355-370.

Mithaug, D. E. (1978). Case study in training generalized instruction-following responses to preposition-noun combinations in a severely retarded young adult. *American Association for the Severely and Profoundly Handicapped Review*, 3, 230-245.

Mithaug, D. E. & Wolfe, M. S. (1976). Employing task arrangements and verbal contingencies to promote verbalizations between retarded children. *Journal of Applied Behavior Analysis*, 9, 301-314.

Murdock, G.Y., Garcia, E.E., & Hardman, M.L. (1977). Generalizing articulation training with trainable mentally retarded subjects. *Journal of Applied Behavior Analysis*, 10, 717-733.

Nutter, D., & Reid, D. H. (1978). Teaching retarded women a clothing selection skill using community norms. *Journal of Applied Behavior Analysis*, 11, 475-487.

Page, T. J., Iwata, B., & Neef, N. A. (1976). Teaching pedestrian skills to retarded persons: Generalization from the classroom to the natural environment. *Journal of Applied Behavior Analysis*, 9, 433-444.

Parsonson, B. S., & Baer, D. M. (1978). Training generalized improvisation of tools by preschool children. *Journal of Applied Behavior Analysis*, 11, 363-380.

Reese, S. C. & Filipczak, J. (1980). Assessment of skill generalization: Measurement across setting, behavior, and time in an educational setting. *Behavior Modification*, 4, 209-223.

Reid, D.H., & Hurlbut, B. (1977). Teaching nonvocal communication skills to multihandicapped retarded adults. *Journal of Applied Behavior Analysis*, 10, 591-603.

Richman, J. S., Sonderby, T., & Kahn, J. V. (1980). Prerequisite vs. in vivo acquisition of self-feeding skill. *Behavior Research and Therapy, 18*, 327-332.

Risley, R., & Cuvo, A. J. (1980). Training mentally retarded adults to make emergency telephone calls. *Behavior Modification, 4*, 513-526.

Rogers-Warren, A., & Warren, S. F. (1980). Mands for verbalization: Facilitating the display of newly trained language in children. *Behavior Modification, 4*, 361-382.

Russo, D., & Koegel, R. L. (1977). A method for integrating an autistic child into a normal public school classroom. *Journal of Applied Behavior Analysis, 10*, 579-590.

Rychtarick, R. G. & Bornstein, P. H. (1979). Training conversational skills in mentally retarded adults: A multiple baseline analysis. *Mental Retardation, 17*, 289-293.

Schleien, S. J., Ash, T., Kiernan, J., & Wehman, P. (1981). Developing independent cooking skills in a profoundly retarded woman. *Journal of the Association for the Severely Handicapped, 6*, 23-29.

Schriebman, L., & Carr, E. G. (1978). Elimination of echolalic responding to questions through the training of a generalized verbal response. *Journal of Applied Behavior Analysis, 11*, 453-463.

Simic, J., & Bucher, B. (1980). Development of spontaneous manding in language deficient children. *Journal of Applied Behavior Analysis, 13*, 523-528.

Smeege, M. E., Page, T. J., Iwata, B. A., & Ivancie, M. J. (1980). Teaching measurement skills to mentally retarded students: Training, generalization and follow-up. *Education and Training of the Mentally Retarded, 15*, 224-230.

Strain, P. S., Shores, R. E., & Timm, M. A. (1977). Effects of peer social initiations on the behavior of withdrawn preschool children. *Journal of Applied Behavior Analysis, 10*, 289-298.

Tofte-Tipps, S., Mendonca, P., & Peach, R. V. (1982). Training and generalization of social skills: A study with two developmentally handicapped, socially isolated children. *Behavior Modification, 6*, 45-71.

Trap, J. J., Milner-Davis, P., Joseph, S., & Cooper, J. O. (1978). The effects of feedback and consequences on transitional cursive letter formation. *Journal of Applied Behavior Analysis, 11*, 381-394.

Tucker, D. J., & Berry, G. W. (1980). Teaching severely multihandicapped students to put on their own hearing aids. *Journal of Applied Behavior Analysis, 13*, 65-75.

VanBiervliet, A., Spangler, P. F., & Marshall, A. M. (1981). An ecobehavioral examination of a simple strategy for increasing mealtime language in residential facilities. *Journal of Applied Behavior Analysis, 14*, 295-305.

Vogelsburg, R. T., Rusch, F. R. (1979). Training severely handicapped students to cross partially controlled intersections. *Journal of the Association for the Severely Handicapped, 4*, 264-273.

Warren, S. F., Baxter, D. K., Anderson, S. R., Marshall, A., & Baer, D. M. (1981). Generalization of question-asking by severely retarded individuals. *Journal of the Association for the Severely Handicapped, 6*, 15-22.

Wehman, P., Renzaglia, A., Berry, G., Schultz, R., & Karan, O. (1978). Developing a leisure skill repertoire in severely and profoundly handicapped persons. *American Association for the Education of the Severely/Profoundly Handicapped Review, 3*, 162-172.

# THE IMPACT OF FUNCTIONAL TRIAL SEQUENCING UPON GENERALIZATION

**Ann K. Berman and**
**Cheryl L. Opalski**

Parents and teachers are often puzzled and frustrated by children with developmental disabilities who tie their shoes at home yet not at school; who do not brush their teeth at home yet brush their teeth every day at school; who can operate one type of clothes washer but not other types. Such examples represent a failure to obtain skill generalization. Traditionally, many researchers considered skill generalization to be a passive phenomenon (Stokes & Baer, 1977). In other words, the transfer of a skill from a training to a nontraining situation was considered to be automatic; no specific training techniques were thought to be necessary to promote this transfer.

However, it is now broadly recognized that the manner in which a skill is initially taught may contribute significantly to successful skill generalization. A feature of training which has received considerable attention in relation to skill acquisition is the manner in which the training trials are sequenced. It is the purpose of this review to examine the influence of various methods of trial sequencing on the generalization process. Specifically, 34 studies were reviewed which addressed the issue of skill generalization and/or trial sequencing in the moderately to profoundly handicapped. Special attention was paid to the type of trial sequencing used by experimenters and how that sequencing might affect skill generalization.

## Definitions

Research into trial sequencing requires establishing clear definitions of several concepts. ***Trial sequencing*** is the process of presenting a stimulus across time and situation. Trial sequencing is considered ***functional*** if it results in a response which is trusted to operate in natural environments (Stokes and Baer, 1977). Thus, the mechanism for programming generalization is called ***functional trial sequencing***.

This review explores the impact of the four methods of trial sequencing upon generalization. Since there is variation in how trial sequencing is defined from one study to the next, the following definitions of trial sequencing types are used in this review:

(1) Massed Trials: Trials which occur so closely together that no other behavior is emitted between trials (Mulligan, Lacy, & Guess, 1982).

(2) Distributed Trials: Trials which occur in such a way that trials from another program occur between two trials from the same program (Mulligan, et al., 1982).

(3) Serial Trials: Trials in which one item was trained to criterion, a second item was trained to criterion, and a third item was trained to criterion, etc. (Waldo, Guess, & Flanagan, 1982).

(4) Concurrent Trials: Trials which involve training several items simultaneously to a specified level of criterion (Waldo, et al., 1982).

Having established these definitions, it should be noted that methods of sequencing are generally combined to form four basic types:

(1) *massed-serial*,
(2) *massed-concurrent*,
(3) *distributed-serial*, and
(4) *distributed-concurrent*.

**Variables**

The current literature suggests that trial sequencing, in and of itself, does not significantly impact skill generalization. However, three variables were identified which, when combined with specific types of trial sequencing, do affect the acquisition and generalization of a skill. These variables are:

(1) the functionality of the skill,
(2) the purchase power of the skills, and
(3) the environmental relevance of the skill.

No one factor was found to be solely responsible for success or failure of generalization. Each of these three variables is addressed separately in this review.

Both the method of sequencing and the variables within sequences, as they relate to generalization, have been examined for this review. It should be noted that the studies reviewed varied along a number of important dimensions including age of subjects (i.e., infant to adult), the type of skill being taught (e.g., use of nonsense syllables, riding a bus, etc.), the degree to which the generalization environment differed from the training environment, and the implied meaning of the term "generalization." Conclusions drawn on the basis of this review are, therefore, of a general and tentative nature.

## The Functionality of the Skill

The literature indicates that functional skills generalize more successfully than nonfunctional skills (see Table 2-1). A skill is considered functional if it has utility for the subjects in their natural environments. According to the definitions used in this review, examples of functional skills include: boarding a public bus (Coon, Vogelsberg, & Williams, 1981), playing pinball (Hill, Wehman, & Horst, 1982), asking curious questions (Hung, 1977), coin counting (Lowe & Cuvo, 1976), and cleaning a restroom (Cuvo, Leaf, & Borakove, 1978). Nonfunctional skill examples include: touching head on cue (Koegel & Rincover, 1977), receptive labeling skills (Charlop, 1983), and some verbal skills (Garcia, 1974; McLean & McLean, 1974; Stevens-Long, Schwarz, & Bliss, 1976; and Guess & Baer, 1973).

Skills taught in a distributed format tend to be functional whereas about 50% of those skills taught in a massed format are rated nonfunctional (see Table 2-2). Thus, generalization tends to be more successful with functional skills, and functional skills tend to be related to distributed rather than massed sequencing.

Three factors within the general issue of functionality seem relevant: (1) whether the skill is initiative or receptive in nature, (2) whether the skill employs cues and/or stimuli already present in the individual's natural environments, and (3) whether the skill uses a form of the behavior already within the individual's repertoire.

### Initiative or Receptive Nature of Task

The individual controls where and when to use an *initiative skill*. On the other hand, a *receptive skill* is one that always

Table 2-1

**Breakdown of Generalization Results According to the Functionality of the Skill**

| | Studies addressing functional skills | Studies addressing nonfunctional skills |
|---|---|---|
| Number of skills[1] | 18 | 10 |
| Studies reporting at least some generalization | 18 (100%) | 9 (90%) |
| Studies reporting 100% generalization immediately | 14 (78%) | 0 (0%) |
| Studies reporting 100% generalization after further training | 3 (17%) | 3 (30%) |
| Studies reporting less than 100% generalization even after further training | 1 (5%) | 6 (60%) |

Table 2-2

**The Relationship Between Functionality and Trial Sequencing in the Studies Reviewed**

| | Number of Studies[1] | Number Functional | Percent Functional |
|---|---|---|---|
| Massed-Serial | 15 | 7 | 47% |
| Massed-Concurrent | 4 | 2 | 50% |
| Distributed-Serial | 1 | 1 | 100% |
| Distributed-Concurrent | 8 | 8 | 100% |
| Total | 28 | 18 | 64% |

1. These include only the studies reviewed which addressed generalization using one type of trial sequencing.

depends on cues from another individual in order for the behavior to occur appropriately. The majority of skills trained in a massed-serial and massed-concurrent sequence are receptive in nature while those skills taught in a distributed-serial or distributed-concurrent format tend to be initiative (see Table 2-3).

Skills which require independence or initiation of a behavior on the part of the child are characteristically more functional in the child's natural environment, while receptive skills are less useful in any environment outside of the training situation. For example, passively responding to an experimenter's three questions about cue cards (Clark & Sherman, 1975) requires receptive skills, while a self-help skill such as washing hands (Zifferblatt, Burton, Horner, & White, 1977) requires initiative. Overall, skills requiring initiative on the part of the child generalize more successfully than receptive skills (see Table 2-4). However, the results are not conclusive since the initiative/receptive nature of the task was not an experimental variable within the studies reviewed.

**Use of Cues and Stimuli within the Natural Environment**

Coon, et al. (1981) trained one retarded person to board and ride a public bus by providing instruction in the classroom and in public places. The skill did not successfully generalize until natural environment training was implemented. Marholin, O'Toole, Touchette, Berger, and Doyle (1979) successfully trained four retarded people to ride a bus to a specified destination, order an item, and purchase the item by training in the classroom, on a bus, and in a shopping center. Generalization on these skills was successful. Coon et al. used a massed-serial format while Marholin et al. used a distributed-concurrent format. Thus, successful generalization can not be attributed to the trial sequencing. Rather, the results suggest that generalization may have been facilitated by incorporating into training the concrete, dependable stimuli present in the natural environment for the skill.

Some studies probed for skill generalization in environments in which the child did not ordinarily operate (i.e., a wooded area with a stranger, Koegel & Rincover, 1977; Rincover & Koegel, 1975). In these studies, the children were placed in a nonnatural setting and asked to perform a nonfunctional task. It is interesting to note that in some cases, a portion of the sample did

Table 2-3

**Relationship between Trial Sequencing and Receptive/Initiative Nature of the Skill**

| | Number of Studies | Studies Addressing Receptive Skills | Studies Addressing Initiative Skills |
|---|---|---|---|
| Massed-Serial | 15 | 12 (80%) | 3 (20%) |
| Massed-Concurrent | 4 | 3 (75%) | 1 (25%) |
| Distributed-Serial | 1 | 0 (0%) | 1 (100%) |
| Distributed-Concurrent | 8 | 1 (12.5%) | 7 (87.5%) |

Table 2-4

**Breakdown of Generalization Results According to the Initiative/Receptive Nature of the Skill**

| | Studies addressing receptive skills | Studies addressing initiative skills |
|---|---|---|
| Number of studies | 16 | 12 |
| Percent functional | 5 (31%) | 12 (100%) |
| Studies reporting at least some generalization | 15 (94%) | 12 (100%) |
| Studies reporting 100% generalization immediately | 5 (33%) | 8 (67%) |
| Studies reporting 100% generalization after further training | 3 (19%) | 3 (25%) |
| Studies reporting less than 100% generalization even after further training | 7 (44%) | 1 (8%) |

respond as trained. However, the ability of an individual to generalize a nonfunctional response such as "head-tapping to a stranger in the woods" is of questionable value in that person's life.

Another aspect of the Rincover and Koegel study (1975) is that the children who failed to generalize were further trained, and an attempt was made to discern the variables to which the children attended. The subjects who failed to generalize responded to the therapist raising his hand, the therapist letting go of the subject's hand, the initial prompt, and the furniture placement. The children were overselective, attending to physical movements or objects rather than to the verbal commands or the people themselves. Thus, when the children were prompted in a strange environment, they called upon past experience to cue their response. These results suggest that natural environment training would be advantageous since many relevant cues and stimuli would be present in both training and nontraining situations.

### Use of Behaviors within the Subject's Repertoire

Using forms of behavior already within the subject's repertoire has implications for generalization. In a study by Charlop (1983), three echolalic children generalized while three mute children did not. Significant to these results is that part of the task was to echo the experimenter prior to handing the experimenter an object. Perhaps the use of a behavior already in the child's repertoire facilitated the generalization process. When the child was not allowed to echo, performance dropped considerably. Charlop suggests that the incorporation of verbal response may have facilitated acquisition and generalization for the echolalic children in that they could more actively participate in the task. The mute children had no such opportunity.

## The Purchase Power of the Skill

After functionality, the next issue which surfaces is the purchase power of the behaviors, especially when failure to generalize is reported. Satiation and diminished motivation are factors related to trial sequencing which seem to impact successful generalization. A massed sequence requires many highly repetitious trials which might prompt boredom with the task and reinforcer situation, whereas a distributed sequence involves less

repetition and the opportunity to perform other behaviors between trials.

**Satiation**

Massed trial sequences require highly repetitious responding by the children. Reinforcement (food or verbal praise) is most often offered on a continuous schedule (Rincover & Koegel, 1975) or on a ratio schedule, usually a variable ratio schedule of 3:1 (Charlop, 1983; Simic & Bucher, 1980). Two to 20 correct trials were required prior to the generalization probes. Charlop (1983) as well as Rincover and Koegel (1975) reported less than 100% generalization. Furthermore, Simic and Bucher (1980) reported no generalization even though reinforcement was offered during generalization probe sessions. Perhaps the children grew tired of the offered reinforcement.

However, massed trial sequences do not always lead to satiation and problems with generalization. Panyan and Hall (1978) required two retarded girls to reach criterion or complete 200 trials. Waldo, Guess, and Flanagan (1982) required six correct responses for each of three receptive labels before moving to another set. However, movement to another portion of the set was allowed when criterion was achieved on one portion. Indeed, the potential for overlearning and insufficient reward was a high risk. In both of these studies, generalization did occur after massed-concurrent trials, but was less successful after massed-serial trials. Thus, the massed nature of the trials, with continuous reinforcement did not always lead to satiation. Rather, another feature within the trial sequencing must have been a significant variable in the generalization results.

While overlearning may facilitate fluency (White & Haring, 1980), insufficient payoff risks extinction of the skill. Ability to perform the task may not have been represented by performance if the individual is satiated on the offered consequences. The law of diminishing returns is likely to be applicable with respect to the motivational power of a piece of candy to evoke a behavior when the child has already received many pieces of candy within a short period of time.

**Motivation**

An issue related to satiation is the motivation of the individual to maintain the skill. Some reinforcers may be less susceptible to

satiation than others and thus maintain the individual's motivation to perform the task.

Bateman (1975) applied Premack's principle that, "for any pair of responses, the independently more probable response will reinforce the less probable one" (p. 604) to show that the occupational behavior of two severely retarded people could be modified. Bateman showed that the subjects were willing to perform a less preferred skill to earn time to perform a desired skill. The chance to perform a desired skill was a natural consequence of the subject's behavior and provided sufficient motivation to complete the less preferred skill.

Although the less preferred skills were quite repetitive (winding balls of wool and plug assembly), as were the preferred activities (sewing a bathmat and sealing plastic bags), successful generalization resulted. While food often seemed to risk satiation, the opportunity to work on a preferred task maintained motivation in this case. So, the motivating power of the offered reinforcement was perhaps more determinant of successful generalization than the sequencing employed.

Another issue related to motivation is the presence or absence of reinforcement during generalization probes. If generalization trials are not artificially reinforced and there is no reinforcement intrinsic to the generalization setting, generalization trials tend to promote extinction of the behavior. Koegel and Rincover (1977) addressed this issue. They trained six autistic children to imitate a behavior, modeled by a therapist, when provided with the verbal cue, "Do this." Each correct response was consequated with food. All six children generalized the receptive skill. However, without reinforcement the skill did not transfer to the generalization setting. A noncontingent reinforcement schedule (NCR) was implemented in further maintenance probes. The first NCR recovered correct responses by the children. Further NCR probes failed to recover correct responding or resulted in low responding.

Therefore, while failure to generalize was reported in several studies, the reported data indicate that generalization may have initially occurred but was extinquished for lack of reinforcement. Indeed, if there is no positive consequence for a behavior in the natural environment, the behavior may not generalize or maintain.

## The Environmental Relevance of the Skill

It was hypothesized that the extent to which the training stimuli resembled stimuli in the natural environment was indicative of generalization success or failure. Thus, a skill taught in a natural environment would generalize more successfully in a natural environment (e.g., domestic, vocational, community) than a skill trained in a contrived environment (e.g., therapy, school). Stokes and Baer (1977) suggested that, when training for generalization, the trainer "use stimuli that are likely to be found in the generalization setting in the training setting as well" (p. 364). In consideration of this issue, the studies were reviewed in terms of the impact of the types of environments (natural and contrived) in which training and generalization probes occurred.

The above hypothesis was not fully supported by the studies reviewed (see Table 2-5). The similarity between the training environment and the generalization environment, in and of itself, was not found to be a major factor in generalization results. However, the studies did not thoroughly test this hypothesis for the following reasons:

(1) The sample size was small (e.g., only one study examined training in a natural environment only and probed in a contrived).
(2) Environmental variables within the studies were not necessarily controlled.

Since the results of the studies reviewed neither proved or disclaimed Stokes and Baer's position, two specific factors within the area of environmental relevance, discrimination of stimuli and independence, were examined.

### Discrimination of Stimuli

To meet the demands of change, a child needs the ability to select relevant stimuli from his environment. Here, sequencing during training gained relevance. It was hypothesized that increased opportunities to make appropriate discriminations would facilitate generalization. Schroeder and Baer (1972) hypothesized that there was a greater opportunity to compare in concurrent training than in serial training. In serial training, the subject made one response repetitiously to a fairly constant stimulus or set of stimuli. On the other hand, in concurrent training, several items are trained simultaneously, thus the individual must attend to the discriminating features of each task.

Table 2-5

**The Impact of Training/Probe Environments on Generalization (in Studies Assessing Generalization Across Settings)**

| Training | Probe | Number of Students | Percent that Reported Some Successful Generalization | Percent that Reported 100% Generalization | Percent that Reported 100% After Additional Training | Percent that Reported Less Than 100% Even After Additional Training |
|---|---|---|---|---|---|---|
| Contrived | Natural | 8 | 100% | 50% | 25% | 25% |
| Contrived & Natural | Natural | 2 | 100% | 50% | 50% | 0% |
| Natural | Contrived | 1 | 100% | 100% | 0% | 0% |
| Contrived | Contrived | 3 | 100% | 100% | 0% | 0% |
| Natural | Natural | 2 | 100% | 100% | 0% | 0% |





BEST COPY AVAILABLE

Three of the reviewed studies specifically compared serial and concurrent training and related these sequences to generalization results (Panyan & Hall, 1978; Schroeder & Baer, 1972; Waldo, Guess, & Flanagan, 1982). In all cases, concurrent training proved to be superior to serial training in facilitating generalization.

However, concurrent versus serial training may not have been the only variable influencing the discrimination of stimuli. For example, Lovaas, Koegel, Simmons, and Long (1973) examined generalization across stimuli with 20 autistic children in behavior therapy. Imitative language, meaningful speech, and social and self-help skills were trained concurrently. These behaviors were trained as alternatives to inappropriate behaviors (self-stimulation and echolalia). "To assess generalization, the children were observed in a room separate from, and not associated with, the training situation and in the company of an unfamiliar adult. The room was equipped with toys (wagon, paper and crayons, a bobo doll, etc.)" (p. 139). The child's behavior in the new setting with the new stimuli was observed. In addition, the child's behavior at home was observed. When the children stayed home with their parents who had learned the techniques, they generalized the behavior to that setting. However, the children who returned to a state hospital did not generalize to that setting

Though generalization occurred in some cases, the reason for the occurrence was not clear. Lovaas, et al. (1973) attributed the generalization success to the varied stimuli; however, another interpretation is that parent training facilitated generalization.

**Independence**

It was hypothesized that those skills which allowed a step toward independence in a natural environment would generalize more successfully than skills which left the subject dependent on others in the environment. Performance of a skill was considered independent if a subject could perform or maintain the skill after instructional support was withdrawn. Self-direction and self-initiation were two qualities considered representative of independent behavior.

A distributed sequence was found to be used with behaviors leading to independent functioning more often than a massed one. Distributed sequences tended to involve a set of responses

each naturally cued by the previous response. For example, Marholin et al. (1979) trained four retarded subjects to ride a bus to a fast-food restaurant, order food items, and pay for the items. The subjects successfully acquired these skills and generalized them to a new setting. The subjects were responsible for their own behavior and independent of trainers. These skills allowed the subjects to operate independently in their natural environment.

In contrast, many skills taught in a massed sequence left the subject dependent upon another person. Clark and Sherman (1975) trained three children to respond to questions about pictures. The child did not successfully generalize these skills to new items. Performance of this skill was dependent upon presentation of cards and questioning by an experimenter. It was not until after additional discrimination training that the children generalized the skill. Even so, the skill could not be performed independently.

## Conclusion

The nature of the task, the environment, and the task's impact are factors which influence skill generalization as much, if not more, than the scheduling of the trials. Futhermore, this review of the literature found that the term "skill generalization" is used nebulously, is difficult to define, and is difficult to train and measure.

Generalization is affected by deviation from the natural cues and consequences present in the environment. It is important for a skill to be necessary in a natural environment. So, generalization seems to be affected by time, setting, people, and the nature of the behavior. Stokes and Baer (1977) state: "Perhaps the most dependable of all generalization programming mechanisms is... the transfer of behavioral control from the teacher-experimenter to stable, natural contingencies that can be trusted to operate in the environment to which the subject will return, or already occupies" (p. 353).

Certain trial sequencing formats tend to contain factors which influence the ability to generalize. However, across-the-board conclusions about any specific format could not be drawn without consideration of trial sequence characteristics. Some of the best predictors of success are whether the skills taught are functional, whether they allow for self-initiation and indepen-

dence, and whether they are reinforced in a natural context. From the studies, it was noted that distributed sequences tend to have features important in promoting generalization. Indeed, it may be argued that trial sequencing itself does not affect generalization, but rather that the type of skill trained should determine both the type of trial sequencing used and the generalization results. In the current literature, the impact of trial sequencing upon generalization is not an issue easily separable from perhaps more significant, extra-sequence influences.

## References


Bateman, S. (1975). Application of Premack's generalization on reinforcement to modify occupational behavior in two severely retarded individuals. *American Journal of Mental Deficiency, 79*, 604-610.

Charlop, M. H. (1983). The effects of echolalia on acquisition and generalization of receptive labeling in autistic children. *Journal of Applied Behavior Analysis, 16*, 111-126.

Clark, H. B. & Sherman, J. A. (1975). Teaching generative use of sentence answers to three forms of questions. *Journal of Applied Behavior Analysis, 8*, 321-330.

Coon, M. E., Vogelsberg, R. I., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for Severely Handicapped, 6*, 46-53.

Cuvo, A. J., Leaf, R. B., & Borakove, L. S. (1978). Teaching janitorial skills to the mentally retarded: Acquisition, generalization, and maintenance. *Journal of Applied Behavior Analysis, 11*, 345-355.

Garcia, E. (1974). The training and generalization of a conversational speech form in nonverbal retardates. *Journal of Applied Behavior Analysis, 7*, 137-149.

Guess D. & Baer, D. M. (1973). An analysis of individual differences in generalization between receptive and productive language in retarded children. *Journal of Applied Behavior Analysis, 6*, 311-329.

Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behavior in severely handicapped youth-pinball machine use. *Journal of the Association for Severely Handicapped, 6*, 38-44.

Hung, D. W. (1977). Generalization of "curiosity" questioning behavior in autistic children. *Journal of Behavior Therapy and Experimental Psychiatry, 8*, 237-245.

Koegel, R. L. & Rincover, A. (1977). Research on the differences between generalization and maintenance in extra-therapy responding. *Journal of Applied Behavior Analysis, 10*, 1-12.

Lovaas, O. I., Koegel, R., Simmons, J. Q., & Long, J. S. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis, 6*, 131-166.

Lowe, M. L. & Cuvo, A. J. (1976). Teaching coin summation to the mentally retarded. *Journal of Applied Behavior Analysis, 9*, 483-489.

Marholin, D., O'Toole, K. M., Touchette, P. E., Berger, P. L., & Doyle, D. A. (1979). "I'll have a big mac, french fries, large coke, and apple pie,"...or teaching adaptive community skills. *Behavior Therapy, 10*, 236-248.

McLean, L. P. & McLean, J. E. (1974). A language training program for nonverbal autistic children. *Journal of Speech and Hearing Disorders, 39*, 186-193.

Mulligan, M., Guess, D., Holvoet, J., & Brown, F. (1980). The individualized curriculum sequencing model (1): Implications from research on massed, distributed, or spaced trial training. *Journal of the Association for the Severely Handicapped, 5*, 325-336.

Mulligan, M., Lacy, L., & Guess, D. (1982). Effects of massed, distributed, and spaced trial sequencing on severely handicapped students' performance. *Journal of the Association for the Severely Handicapped, 7*, 48-61.

Panyan, M. C. & Hall, V. (1978). Effects of serial versus concurrent task sequencing on acquisition, maintenance, and generalization. *Journal of Applied Behavior Analysis, 11*, 67-74.

Rincover, A. & Koegel, R. L. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis, 8*, 235-246.

Schroeder, G. L. & Baer, D. M. (1972). Effects of concurrent and serial training on generalized vocal imitation in retarded children. *Developmental Psychology, 6*, 293-301.

Simic, J. & Bucher, B. (1980). Development of spontaneous manding in language deficient children. *Journal of Applied Behavior Analysis*, 13, 523-528.

Stevens-Long, J., Schwarz, J. L., & Bliss, D. (1976). The acquisition and generalization of compound sentence structure in an autistic child. *Behavior Therapy*, 7, 397-404.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10, 349-367.

Waldo, L., Guess, D., & Flanagan, B. (1982). Effects of concurrent and serial training on receptive labeling by severely retarded individuals. *Journal of the Association for the Severely Handicapped*, 6, 56-65.

White, O.R., & Haring, N.G. (1980). *Exceptional Teaching*. Columbus, Ohio: Bell and Howell Company.

Zifferblatt, S. M., Burton, S. D., Horner, R., & White, T. (1977). Establishing generalization effects among autistic children. *Journal of Autism and Childhood Schizophrenia*, 7, 337-347.

# SELF-MONITORING AND SKILL GENERALIZATION: A REVIEW OF CURRENT RESEARCH

**Kathleen A. Liberty**

When settings, managers, contingencies, consequences, cues, or prompts change, students often fail to display acquired behaviors. This failure to perform under circumstances which differ from training has been called a failure to generalize, and constitutes a serious threat to the functional impact of instructional programs. Consequently, current research efforts are directed at identifying practical solutions to the problem. Much of the current research effort in generalization is directed at modifying the training environment and/or instructional procedures to facilitate generalization. Another approach is to teach the handicapped person to control the variables that seem to affect generalization. In most training situations, antecedents (cues and prompts) and consequences (e.g., reinforcers) are controlled by teachers, therapists, and other caregivers. If the handicapped person can control those variables herself, the differences between settings may be minimized, and generalization facilitated. Self-control techniques, in which the individual controls these variables, offer hope for facilitating generalization.

The three most common self-control techniques include self-instruction (control of antecedents), self-reinforcement, and self-monitoring. Self-monitoring may be defined as the act of counting or otherwise registering overtly the occurrence of the target behavior. The act of self-monitoring implies that the individual perform an observable behavior which functionally identifies the completion (or initiation) of a particular behavior. Such signaling is an integral component of antecedent and consequent control as well, since self-instruction implies that the individual determine when one behavior has been completed and another is to be initiated, and since self-reinforcement implies that the individual determine when a behavior has been completed (Kazdin, 1980). Therefore, the understanding and examination of self-monitoring is an essential step in the study of self-control.

## The Act of Self-Monitoring

Self-monitoring involves (at a minimum) two overt behaviors: the target behavior (i.e., the behavior that is self-monitored) and the act of self-monitoring itself. While there is obvious interest in the effects of self-monitoring on the target behavior, the act of self-monitoring itself has not been systematically investigated(Brigham, 1978).

One important set of questions involves around the instrumentation of self-monitoring. Self-monitoring behaviors that have been utilized by subjects in studies include: marking tallies on a piece of paper with a pencil (Jackson & Martin, in press; Broden, Hall, & Mitts, 1971; Gottman & McFall, 1972; Nelson, Lipinski, & Black, 1976; Zegiob, Klukas, & Junginger, 1978); coloring squares on a piece of paper (Nelson, Hay, Devany, & Koslow-Green, 1980); pushing a button on a computer (Brodsky, Lepage, Quiring, & Zeller, 1970; Epstein, Webster, & Miller, 1975); crossing off numbers on a form (Hayes & Cavior, 1980); pressing a telegraph key (Hayes & Nelson, 1983); marking a "+" or a "−" in squares on a form (Horner & Brigham, 1979); marking off squares on a form (Lipinski & Nelson, 1974); actuating a wrist counter (Ollendick, 1981); and actuating a hand-held frequency counter (Willis & Nelson, 1982). Unfortunately, investigations of the relationship between the form of self-monitoring and changes in the target behavior, the form of self-monitoring and the speed of training, the form of self-monitoring and the reliability of self-monitoring, and the form of self-monitoring and the maintenance of self-monitoring have not been reported.

Another aspect of the act of self-monitoring that is of importance is the timing of self-monitoring (e.g., after each behavior, before each behavior, after several behaviors, during an interval of responding). Although timing may influence the effect of self-monitoring (Nelson & Hayes, 1981; Nelson et al., 1980; Albion, 1983), it is often neglected in experimental reports.

Another important variable, relating both to the act of self-monitoring and its effects on the target behavior is the reliability of self-monitoring. Unfortunately, many investigators fail to report reliability. When it is included, reliability is reported as a Pearson product moment correlation, which indicates reliability of scores in terms of level changes, but does not estimate actual accuracy. For example, self-monitoring of 1, 2, and 3 would correlate perfectly ($r = 1.0$) with an observer's scores of 10, 20,

and 30, while accuracy calculated by traditional reliability estimate formulas (e.g., agreements ÷ agreements + disagreements × 100) would be 10%. In addition, reliabilities are presented as means across sessions and/or across subjects; such data do not demonstrate how reliabilities may fluctuate, and make direct comparisons of the day to day changes in both self-monitoring accuracy and the effects of self-monitoring impossible to determine.

The reliability of subjects has ranged from −.06 and .17 (Hayes & Cavior, 1980), to .52 and .86 (Lipinski & Nelson, 1976), to .90 and .95 (Willis & Nelson, 1982). In a series of studies, Nelson, Hay, Devany, and Koslow-Green (1980) reported reliabilities (as Pearson product moment correlations) ranging from .07 to .79 for regular elementary school children in first, third, and fifth grades. Accuracy increased across age for self-monitoring positive verbalizations, and decreased across age for counting inappropriate verbalizations. Reliability for handicapped subjects has been reported as 80% and 87% (Horner & Brigham, 1979) and 96 to 99% (Jackson & Martin, in press). In some studies, however, investigators have reported that subjects who had been trained to self-monitor were not observed to do so after training ceased (Broden, Hall, & Mitts, 1971; Seymour & Stokes, 1976; Gottman & McFall, 1972; Zegiob, Klukas, & Junginer, 1978; James, 1981).

Accuracy of self-monitoring may be improved by both notifying subjects that reliability will be checked (Hayes & Horn, 1982) and by reinforcing accuracy (e.g., Turkewitz, O'Leary, & Ironsmith, 1975; Fixsen, Phillips, & Wolfe, 1972; Risley & Hart, 1968). Unfortunately, specific positive reinforcement for self-monitoring may confound examination of changes in target behavior (e.g., Epstein, Webster, & Miller, 1978; Epstein, Miller, & Webster, 1978), reduce the probability of the generalization of self-monitoring itself (Stokes & Baer, 1977), and/or establish self-monitoring as a behavior which competes with the target behavior (Epstein, Webster, & Miller, 1978). In addition to questions on the relationship between effects on target behaviors and self-monitoring reliability, the issue is of concern because the reliability of self-monitoring is one measurable aspect of the success of self-monitoring training. If, as proposed, self-monitoring is to be used to promote cross-setting skill generalization, it is to be expected that self-monitoring will be trained. Systematic evaluation of the success of training methods will involve measures of the accuracy of self-monitoring itself.

Another focus of concern in studying self-monitoring is how subjects were trained to self-monitor. Most often, researchers have used a combination of verbal directions, demonstrations, practice opportunities, reinforcement, and error correction procedures to teach self-monitoring, and training is usually conducted in a single session (Gardener, Cleas, & Cole, 1983; Hayes & Nelson, 1983; Lipinski & Nelson, 1974; Matson & Earnhart, 1981). Often experimenters report that additional prompting, feedback, and reinforcement are required to establish or maintain reliable self-monitoring of the target behavior following training or in other settings (e.g., Ollendick, 1981). Unfortunately, experimental data on the acquisition, mastery, maintenance, and generalization of the act of self-monitoring are not presented in the experimental literature. Therefore, questions as to how to train the act of self-monitoring to produce reliable self-monitoring, which itself maintains, and which generalizes across settings (and/or across target behaviors), remain unexplored.

## The Impact of Self-Monitoring on Target Behaviors

Self-monitoring was originally introduced into the research literature as a data collection mechanism, and was used in experiments where the target behavior was of a private nature (e.g., marital fights), or where external observers were obtrusive, or when long periods of data collection were required. However, researchers quickly noted that self-monitoring itself produced changes in the target behavior, and self-monitoring became a therapeutic intervention strategy (Kazdin, 1974; Kazdin, 1980). The majority of studies in the area of self-monitoring have been directed at: (1) the effects of self-monitoring (as an intervention strategy) on a target behavior (e.g., Zohn & Bornstein, 1980; Horner & Brigham, 1979); (2) the impact of different variables on the effects produced by self-monitoring (e.g., Kazdin, 1974; Nelson, Hay, Hayany, & Kostelow-Green, 1980); and (3) a comparison of the effects of self-monitoring with other interventions (e.g., Jackson & Martin, in press; Hanel & Martin, 1980; Rosenbaum & Drabman, 1979).

In general, when self-monitoring is used as an intervention strategy, it has resulted in increases when the target behavior was defined as "positive" and an increase was a therapeutic goal, and in decreases when the behavior was defined as "negative." However, such effects have not been shown consistently.

Many variables which might affect the degree of effect self-monitoring has on the target behavior have been studied (Hayes & Horn, 1982, present a review of the literature in the area). Important issues include:

(1) The direction of change desired for the target behavior. In some studies, behaviors with a postive valence (increase desired) show more effects than behaviors with a negative valence (Willis & Nelson, 1982). However, behaviors with a negative valence show higher levels of self-monitoring reliability. In general, behaviors with higher response strength and "social sensitivity" show greater effects (Hayes & Horn, 1982).

(2) Overt or covert observations of subjects engaged in self-monitoring. Overt observations produce more accurate self-monitoring and may produce increased effects over covert observation (Lipinski & Nelson, 1974).

(3) Reliability of self-monitoring has not been associated with effects obtained. The relationship between accurate self-monitoring and changes in the target behavior are not established (Rosenbaum & Drabman, 1979; references).

The effect of self-monitoring on target behaviors has been compared with other interventions, but it is usually applied in combination with either self-reinforcement or self-instruction (Kazdin, 1980). In general, self-control techniques have been as effective as externally-controlled interventions in changing behavior (Rosenbaum & Drabman, 1979, present a review of the literature in this area).

An examination of the methodology of many studies indicates that the training provided in self-monitoring may confound examination of effects on the target behavior (Albion, 1983; Nelson & Hayes, 1981; Kazdin, 1980). For example, in some studies the subject is not told which behavior is the target until baseline data have been collected, and then training in self-monitoring includes statements as to the desirability of changing the behavior in a certain direction (e.g., Rosins & Martin, in press; Hayes & Nelson, 1983; Gardner, Cless, & Cole, 1983). Subsequent changes in the target behavior may then be the result either of the self-monitoring or of the goal statements during training (Albion, 1983).

Another confounding variable is present when the subject is reinforced by the experimenter for self-monitoring (e.g.,

Fixen, Phillips, & Wolf, 1972; Goyos, Michael, & Martin, 1979; Matson & Earnhart, 1981; Shapiro & Klein, 1980); the *reinforcement is then an additional consequence for* performing the behavior (since self-monitoring follows the behavior). Any increase in behavior, therefore, may be a result of the "extra" experimenter-delivered reinforcement and not just self-monitoring (Epstein, Webster, & Miller, 1975; Zegiob, Klukas, & Junginger, 1978; Goyos, et al. 1979; Albion, 1983). Other possible confounding factors include very short experimental phases (e.g., 20 minutes), failure to report reliability, and vague procedural descriptions (Albion, 1983). Despite these problems, however, self-monitoring does seem to produce changes in the behavior which is self-monitored.

Three major theoretical explanations for effects of self-monitoring on target behaviors have been developed:

(1) Self-monitoring provides a feedback loop; the individual observes and records changes in her own behavior and then adjusts the target behavior (Kanfer, 1970).

(2) Self-monitoring acts to clarify the relationship between the behavior and its consequences; behavior change is the result of the increasing correspondence (for the behaver) of the relationship (Rachlin, 1974).

(3) The events surrounding the initial and continuing use of self-monitoring (e.g., training procedures, materials used to self-monitor) function as stimuli which cue environmental consequences, and those consequences result in changes in the target behavior (Nelson & Hayes, 1981).

However, until additional carefully designed studies accummulate, these explanations must remain theoretical.

If self-monitoring is to be used to facilitate cross-setting generalization of the target skill, the impact of self-monitoring on changes in the target behavior are of import. Far more critical, however, is the value of self-monitoring in maintaining and extending behavior changes established in more traditional methods. For example, a (target) behavior may be acquired and mastered through traditional externally-controlled instructional techniques; self-control skills may then be taught to the student to maintain treatment gains or to promote skill generalization. In such cases, the crucial issue is the value of self-monitoring for

maintaining treatment gains during post-treatment periods in nontraining settings.

Recently, investigators have expanded the study of self-monitoring to include the use of self-monitoring to maintain target behaviors and to facilitate cross-setting generalization. Drabman, Spitalick, & O'Leary (1973) taught a combination of self-control skills (including self-monitoring) to a group of disruptive children in an elementary school. The self-control skills produced maintenance of gains in the target behavior and generalized to other periods of the school day. In a replication, Turkewitz, O'Leary, & Ironsmith (1975) produced similar results, but failed to show cross-setting generalization. Holman & Baer (1979) taught six children to count pages of academic work in writing and math. Effects of self-monitoring were measured both during the individual experimental sessions and in the children's regular classroom (cross-setting generalization). Following baseline, the children were taught to count each page completed by moving a bead on a wrist bracelet. A goal was established for the number of pages to be worked, and a white bead inserted on the counter by the experimenter. The child was directed to try to get to the white bead. During the self-monitoring phases, the child was praised for self-monitoring, both in the 1:1 situation and in the classroom, and for reaching the goal (i.e., the white bead). Similar goal setting and reinforcement were not present in baseline, so measurement of the effects of self-monitoring on the target behaviors (i.e., on task, off task, and disruptive behaviors) was confounded, although the authors reported changes in the desired directions in both the experimental and the classroom settings. The act of self-monitoring transferred to the classroom (i.e., generalized), with high levels of accuracy (reliability of self-monitoring during the study was 90% or better).

Stokes and Baer (1977) have reported that one method of facilitating generalization is to establish a response that generates a stimulus that will occur in both training and nontraining situations ("mediate generalization"). Holman & Baer (1979) suggest that the cross-setting transfer of effects on the target behaviors was mediated by "the discriminative properties of the bracelet [i.e., counter], which served to remind subjects (1) to work, (2) to complete a certain amount of work, (3) to measure ongoing progress, and (4) finally to seek reinforcement from the teacher for having done so" (p. 442). Holman and Baer (1977) conducted two follow-up measures, for months 2-5 and months 7-10 following the conclusion of the study, and self-monitoring was durable throughout, although some of the gains made in target

behaviors were not maintained. Increasing attention to self-monitoring itself may be a function of the promise of self-monitoring and other self-control skills for promoting cross-setting skill generalization (Kazdin, 1980; Holman & Baer, 1979; Baer, Holman, Stokes, Fowler, & Rowbury, 1981; Turkewitz, O'Leary, & Ironsmith, 1975; Kurtz & Neisworth, 1976). Further studies of the use of self-monitoring to maintain and generalize skilled behaviors may help to identify not only the effects of such self-control training, but also the durability and means of any gains. Such questions are of immediate importance and value in any evaluation of the potential of self-control techniques for promoting skill generalization.

## Self-Monitoring by Severely Handicapped Individuals

Severely handicapped subjects have rarely participated in investigations of self-monitoring, although self-monitoring has been used as an intervention agent in studies with moderately and mildly retarded subjects (e.g., Gardner, Clees & Cole, 1983; Nelson, Lipinski, & Black, 1976; Zohn & Bornstein, 1980; Zegiob et al., 1983; Shapiro & Klein, 1980). The effects of self-monitoring as an intervention agent with mildly and moderately retarded workers have replicated results with nonhandicapped subjects. Unfortunately, confounding effects of externally-controlled reinforcement and training procedures, as well as procedural questions of training, instrumentation, and reliability of self-monitoring, limit the applicability of such studies to questions of major importance in the study of self-monitoring as a method of facilitating skill generalization.

In addition, studies have failed to document the reliability of self-monitoring, the efficiency of the training methods used, and whether the act of self-monitoring itself has maintained or generalized. Although the act of self-monitoring itself may be treated like other behaviors trained, specific problems in the method of training used may not only confound effects of self-monitoring on target behaviors (e.g., if self-monitoring is reinforced), but inhibit generalization and maintenance of self monitoring with other behaviors. In addition, the act of self-monitoring can not be trained in isolation. The subject must be trained to count something. If the target skill is also being instructed or intervened on, instruction in self-monitoring may interrupt that process. Using traditional instructional

procedures for both behaviors may then result in confusion, or in problems with both the target behavior and the instructed response of self-monitoring. Finally, specific studies relating to the use of self-monitoring to facilitate cross-setting generalization of target skills have not been reported. The promise of self control strategies for facilitating generalization makes it necessary to investigate with greater precision how to most effectively train self-monitoring, the maintenance of self-monitoring, and the transfer of self-monitoring across settings and across behaviors, as well as to continue to study the effect of self-monitoring on behavior and the transfer of those effects to other settings.

A series of studies has been conducted to probe specific questions relating to the use of self-monitoring to facilitate skill generalization with severely handicapped individuals (see Liberty & Paeth, 1983a; Liberty & Paeth, 1983b; Liberty, 1983).

The purpose of the first study was to examine the acquisition of self-monitoring by a severely handicapped student through an avoidance training procedure, which eliminated the confounding effects of other training procedures while permitting simple integration into ongoing instruction. Measures of the independent use of self-monitoring and the reliability of self-monitoring were used as indices of acquisition. In addition, the effects of self-monitoring on the target behavior were examined.

The subject was 19 years and five months old. He attended a secondary program for severely handicapped youth, located in a public high school. The experimental design consisted of a baseline phase, two training phases, and a contingency change phase. The baseline provided a measure of the rate of production by the subject. During the training phases, the subject was taught to actuate a counter placed on the table next to his work. The training procedure consisted of interrupting work (which delayed reinforcement available for work) for nonactuation of the counter, which was then prompted by gestures (no verbal directions given). Delay of reinforcement was avoided by self-monitoring. No specific reinforcement for self-monitoring was provided. In the second training phase, the subject wore the counter on his wrist. In the last phase, a general contingency for behavior was introduced.

The avoidance training procedure produced rapid acceleration of unprompted and very reliable self-monitoring. Self-monitoring itself was positively reinforced, presumably by the audible click of the counter and the display roll-over. Self-moni-

toring produced changes in the target behavior (rate of production) indicating that self-monitoring acted as a (self-controlled) positive reinforcer sufficient to maintain performance of the target behavior. A similar effect was shown by Nelson et al. (1980), who found that first and fourth graders recorded higher frequencies of the target behavior when the behavior of self-monitoring was coloring squares with a crayon. The authors speculated that the subjects engaged in higher frequencies of the target behavior in order to color squares (self-monitoring), and colored "extra" squares as well. Holman and Baer (1979) also found that subjects prompted teachers to praise them for self-monitoring when they reached the white bead.

Investigators reporting studies in which self-monitoring alone produced increases in target behaviors have speculated that such increases may be attributed to (1) the nature of the training, (2) additional, externally-controlled, reinforcement for self-monitoring, (3) motivational properties of self-monitoring produced by covert self-evaluation, or (4) that all of the events surrounding the act of self-monitoring and training in self-monitoring function as stimuli that cue envionmental consequences that, in turn, effect behavior change (Nelson & Hayes, 1981). In the current study, the training procedures for self-monitoring excluded any statements about desired changes in the target behavior and avoided any additional externally-controlled reinforcement. It is also unlikely that the subject, who was unable to count or recognize numerals, engaged in covert comparisons of the numeral counter displays from session to session (self-evaluation). As an alternative, it may be hypothesized that the behavior of pushing the counter, accompanied by an audible "click" and a changing display, acted to reinforce not only the behavior of self-monitoring, but the target behavior as well.

The purposes of the second study were to:

(1) Examine the maintenance of self-monitoring skills.
(2) Determine if self-monitoring generalized within stimulus classes and across responses, and if so, the extent and nature of such generalization.
(3) Examine the effect of self-monitoring on the target behaviors.

The same subject participated. The first study had concluded 343 days prior to the start of this study; during the intervening period, the subject had not worked on production or self-monitored.

This study incorporated a multiple baseline design across two behaviors, one performed in the classroom (production) and one in a cafeteria work setting (assembly of sack lunches). The phases for production were Baseline I, Self-monitoring I, Baseline II, and Self-monitoring II. The phases for bagging sack lunches were Baseline I, Opportunity to self-monitor (introduced simultaneously with Self-monitoring I of production), Train to self-monitor, and Baseline II.

Production did not maintain at a commensurate level during the 343 intervening days. Self-monitoring did, however, maintain at high levels of reliability and independence. Production was consequated by instructor-controlled reinforcement, while self-monitoring was consequated by subject-controlled reinforcement (counter click and display roll-over). Behaviors that are consequated by self-controlled reinforcement are more resistant to extinction than are behaviors consequated by externally-controlled reinforcement (Weiner & Dubanoski, 1973), and perhaps that is why self-monitoring maintained at higher levels than did production.

The subject did not actuate the counter, even when he had the opportunity to do so, and therefore training in self-monitoring was introduced, and accurate self-monitoring acquired in four sessions. The independent and reliable self-monitoring of production did not transfer to self-monitoring of bagging. In this study, the two behaviors were not of the same response class, and were performed under conditions (e.g., supervision, stimulus materials, setting, time of day) totally unlike one another. Transfer may have occurred if the settings were more alike; for example, if the subject had counted bags in his classroom, with the same supervisor present.

However, wearing the wrist counter (but not actuating it) produced an increase in bagging rate that was of practical significance to the subject (from 46% to 67.5% of normal, with a high rate of 114% of the normal rate), and high rates continued as self-monitoring was trained. Once the opportunity to self-monitor was removed from production, bagging rate dropped. When the opportunity to self-monitor bagging was withdrawn, bagging began decelerating.

Similar effects have been reported in other studies (Broden, et al., 1971; Gottman & McFall, 1972; Zegiob, et al., 1978). There are several hypotheses for the effects of self-monitoring when the act of self-monitoring is itself not observed. The wrist

counter could serve as an $S^D$ for faster work, or it may act as a reinforcer by mediating differences between settings. Stokes and Baer (1977) have reported that one method of facilitating generalization is to establish a response (in this study, self-monitoring), that generates a stimulus (the counter) that will occur in both training and nontraining settings. In the present study, externally-controlled reinforcement was more frequent during production (VI 1.5' vs. VI 5') and any self-controlled reinforcement exerted by the act of self-monitoring was also more frequent during production, since pages were finished more quickly than sacks. These data support the use of self-monitoring to mediate across setting generalization. The counter, which was worn during more frequent reinforcement (during production), may have acquired generalized reinforcing properties that mediated generalization of faster work rates across settings which differed in reinforcement schedules.

The purpose of the third study was to extend and replicate results with different behaviors and with a different subject. Specific issues examined included:

(1) The effects of wearing a wrist counter on two behaviors, one instructed and the other uninstructed, prior to any training in the use of a counter, and whether effects produced in the training setting would transfer to the probe setting.

(2) The effects on the instructed and uninstructed behaviors of training self-monitoring of the instructed behavior, in both the training setting and the probe setting.

(3) Whether self-monitoring would transfer from the behavior on which it was trained (i.e., the instructed behavior) to another behavior (i.e., the uninstructed behavior).

(4) Whether self-monitoring would transfer from the setting in which it was trained to the probe setting.

The subject of this study was an eleven-year old girl (I.Q., 30), attending a special school for handicapped children. The subject's regular classroom served as the training setting. From one to four times per week, the subject attended another classroom in the school for thirty minutes each time. This classroom served as the probe setting. Two target behaviors were measured in both settings: two-word responses to questions, and two-word initiations. Two-word responses were differentially reinforced in the training setting on a fixed ratio of 1; while in the probe setting, they were reinforced on a variable schedule of about 1 to

6. two-word initiations were reinforced at approximately the same ratio (VR 3) in both settings.

Following baseline observations, sessions in the training setting were divided into two approximately equal periods. The subject wore the counter during one of the two periods each day. In the third phase, baseline conditions were reinstituted. In the fourth phase, the subject was trained to self-monitor two-word answers, using an avoidance training procedure. No training in self-monitoring was provided in the probe setting. Instructors in that setting were asked to ignore the wrist counter. In the fifth phase, baseline conditions were reinstated.

Wearing the wrist counter without any training in self-monitoring did not affect performance. Once self-monitoring was trained, however, performance was affected dramatically. The act of self-monitoring generalized across behaviors and across settings for both behaviors. Self-monitoring improved cross-setting generalization of two-word answers from 19% to 49%; performance deteriorated quickly once the opportunity to self-monitor in the probe setting was withdrawn. Self-monitoring did not affect cross-setting performance of two-word initiations. Preliminary analyses of the results suggest that self-monitoring successfully mediated the differences in contingencies between probe and training setting; when there were no differences, self-monitoring did not affect performance. Further, the effect is directly linked to the self-monitoring training, during which the counter acquired discriminative properties, which functioned as self-controlled reinforcement in the probe setting.

The results of these studies suggest first that self-monitoring may effect a change in a target behavior in the training setting, under some circumstances, with severely handicapped subjects. Effects are not dependent on the accuracy of self-monitoring. These data support the findings with mildly and nonhandicapped subjects. These studies also indicate that the avoidance training procedure can produce rapid acquisition of accurate self-monitoring, which is maintained, and which can generalize within response classes, but not across response classes, and which can generalize across settings. Third, these studies support the use of self-monitoring to mediate skill generalization when the reinforcement contingencies differ across settings. Although these studies are limited to two subjects, results are similar to those reported with other groups of subjects, but expand especially on the information on the use of self-monitoring to facilitate skill generalization. Future research with

additional severely handicapped subjects will be conducted to test the results and hypotheses.

## References


Albion, F. (1983). A methodological analysis of self-control in applied settings. *Behavioral Disorders, 8*(2), 87-102.

Baer, D., Holman, J., Stokes, T., Fowler, S., & Rowbury, T. (1981). Uses of self-control techniques in programming generalization. In S. Bijou and R. Ruiz (Eds), *Behavior Modification: Contributions to Education*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Brigham, T. (1978). Self-Control. In C. Cantonia & T. Brigham (eds.), *Handbook of applied behavior analysis*. New York: Irvington Publishers, 259-274.

Broden, M., Hall, R., & Mitts, B. (1971). The effect of self-recording on the class-room behavior of two eighth-grade students. *Journal of Applied Behavior Analysis, 4*, 191-199.

Brodsky, G., LePage, T., Quiring, J., & Zeller, R. (1970). Self-evaluative responses in adolescent retardates. *American Journal of Mental Deficiency, 74*(6), 792-795.

Drabman, R., Spitalick, R., & O'Leary, D. (1973). Teaching self-control to disruptive children. *Journal of Abnormal Psychology, 82*(1), 10-16.

Epstein, L., Miller, P., & Webster, J. (1976). The effects of reinforcing concurrent behavior on self-monitoring. *Behavior Therapy, 7*, 89-95.

Epstein, L., Webster, J., & Miller, P. (1975). Accuracy and controlling effects of self-monitoring as a function of concurrent responding and reinforcement. *Behavior Therapy, 6*, 654-666.

Fixen, D., Phillips, E., & Wolf, M. (1972). Achievement place: The reliability of self-reporting and peer reporting and their effects on behavior. *Journal of Applied Behavior Analysis, 5*, 19-30.

Gardner, W., Cless, T., & Cole, C. (1983). Self-management of disruptive verbal ruminations by a mentally retarded adult. *Applied Research in Mental Retardation, 4*(1), 41-58.

Gottman, J. & McFall, R. (1972). Self-monitoring effects in a program for potential high school dropouts: A time series analysis. *Journal of Consulting and Clinical Psychology, 39*, 273-282.

Goyos, A., Michael, J., & Martin, G. (1979). Self-recording training to teach retarded adults to reinforce work behaviors of retarded clients. *Rehabilitation Psychology, 26*(4), 215-227.

Hanel, F., & Martin, G. (1980). Self-monitoring, self-administration of token reinforcement, and goal-setting to improve work rates with retarded clients. *International Journal of Rehabilitation Research, 3*, 505-517.

Hayes, S., & Cavior, N. (1980). Multiple tracking and the reactivity of self-monitoring: II. Positive behaviors. *Behavioral Assessment, 2*, 283-296.

Hayes, S. & Nelson, R. (1983). Similar reactivity produced by external cues and self-monitoring. *Behavior Modification, 7*(2), 183-196.

Holman, J., & Baer, D. (1979). Facilitating generalization of on-task through self-monitoring of academic tasks. *Journal of Autism and Developmental Disorders, 9*(4), 429-445.

Horner, R., & Brigham, T. (1979). The effects of self-monitoring procedures on the study behavior of two retarded children. *Education and Training of the Mentally Retarded, 14*, 18-24.

Jackson, D. & Martin, G. (in press). Additive effects of components of a self-control package for improving the production of mentally handicapped workers. *Journal of Practical Approaches to Developmental Handicap*.

James, T. (1981). Self-monitoring of stuttering: Reactivity and accuracy. *Behavior Research and Therapy, 19*, 291-296.

Kanfer, F. (1970). Self-monitoring: Methodological limitations and clinical applications. *Journal of Consulting and Clinical Psychology, 35*, 148-152.

Kazdin, A. (1980). Reactive self-monitoring: The effects of response desirability goal setting and feedback. *Journal of Consulting and Clinical Psychology, 42*, 704-716.

Kazdin, A. (1980). *Behavior modification in applied settings*. Homewood, IL: The Dorsey Press.

Kurtz, P., & Neisworth, J. (1976). Self-control possibilities for exceptional children. *Exceptional Children, 42*, 212-217.

Liberty, K., & Paeth, M. (1983). Self-monitoring I: The acquisition of self-monitoring and its effect on the production rate of a severely handicapped adolescent. *Institute for education of severely handicapped children: Washington Research Organization: Annual report FY 82-83.* (U.S. Department of Education, Contract No. 300-82-0634.) Seattle: University of Washington, Experimental Education Unit.

Liberty, K., & Paeth, M. (1983). Self-monitoring II: Maintenance and generalization of self-monitoring and its effects on two target behaviors. *Institute for education of severely handicapped children: Washington Research Organization: Annual report FY 82-83.* (U.S. Department of Education, Contract No. 300-82-0634.) Seattle: University of Washington, Experimental Education Unit.

Liberty, K. (1983). Self-monitoring III: Effects of self-monitoring training on expressive communication: Mediation across settings. *Institute for education of severely handicapped children: Washington Research Organization: Annual report FY 82-83.* (U.S. Department of Education, Contract No. 300-82-0634.) Seattle: University of Washington, Experimental Education Unit.

Lipinski, D., & Nelson, R. (1974). The reactivity and unreliability of self-monitoring. *Journal of Consulting and Clinical Psychology*, 42, 118-123.

Mattson, J., & Earnhart, T. (1981). Programming treatment effects to the natural environment: A procedure for training institutionalized retarded adults. *Behavior Modification*, 5, 27-37.

Nelson, R., Hay, L., Devany, J., & Koslow-Green, L. (1980). The reactivity and accuracy of children's self-monitoring: Three experiments. *Child Behavior Therapy*, 2(3), 1-24.

Nelson, R., & Hayes, S. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification*, 5, 3-14.

Nelson, R., Lipinski, D., & Black, J. (1976). The reactivity of adult retardates' self-monitoring: A comparison among behaviors of different valences, and a comparison with token reinforcement. *The Psychological Record*, 26, 189-201.

Ollendick, T. (1981). Self-monitoring and self-administered overcorrection: The modification of nervous tics in children. *Behavior Modification*, 5(1), 75-84.

Rachlin, H. (1974). Self-control. *Behaviorism*, 2, 94-107.

Risley, T., & Hart, B. (1968). Developing correspondence between nonverbal behavior of pre-school children. *Journal of Applied Behavior Analysis*, 1, 267-281.

Rosenbaum, M., & Drabman, R. (1979). Self-control training in the classroom: A review and critique. *Journal of Applied Behavior Analysis, 12*, 467-485.

Rosine, L., & Martin, G. (in press). Self-management training to decrease undesirable behavior of mentally handicapped adults. *Rehabilitation Psychology*.

Seymour, F., & Stokes, T. (1976). Self-recording in training girls to increase work and evoke staff praise in an institution for offenders. *Journal of Applied Behavior Analysis, 9*, 41-54.

Shapiro, E., & Klein, R. (1980). Self-management of classroom behavior with retarded/disturbed children. *Behavior Modification, 4*(1), 83-94.

Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Turkewitz, H., O'Leary, K., & Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. *Journal of Consulting and Clincial Psychology, 43*(4), 577-583.

Weiner, H., & Dubanoski, R. (1973). Resistance to extinction as a function of self- or externally determined schedules of reinforcement. *Journal of Personality and Social Psychology, 31*, 905-910.

Willis, S., & Nelson, R. (1982). The effects of valence and nature of target behavior on the accuracy and reactivity of self-monitoring. *Behavioral Assessment, 4*, 401-412.

Zegiob, L., Klukas, N., & Junginger, J. (1978). Reactivity of self-monitoring procedures with retarded adolescents. *American Journal of Mental Deficiency, 83*, 156-163.

Zohn, C., & Bornstein, P. (1980). Self-monitoring of work performance with mentally retarded adults: Effects upon work productivity, work quality, and on-task behavior. *Mental Retardation, 18*, 19-25.